問2.1

$$S(\beta_0, \beta_1) \equiv \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \{y_i - (\beta_0 + \beta_1 x_i)\}^2$$

$$\frac{\partial S(\beta_0, \beta_1)}{\partial \beta_0} = -2 \sum_{i=1}^{n} \underbrace{\{y_i - (\beta_0 + \beta_1 x_i)\}}_{= 0}$$

$$= 0$$

$$\therefore \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \beta_0 + \beta_1 x_i$$

$$= n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i$$

$$\frac{\partial S(\beta_0, \beta_1)}{\partial \beta_1} = -2 \sum_{i=1}^{n} x_i \underbrace{\{y_i - (\beta_0 + \beta_1 x)\}}_{= 0}$$

$$\therefore \sum_{i=1}^{n} x_i y_i = \beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2$$

問2.2

$S(\boldsymbol{\beta}) = \sum_i \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\boldsymbol{y} - X\boldsymbol{\beta})^T (\boldsymbol{y} - X\boldsymbol{\beta})$ を$\boldsymbol{\beta}$で偏微分すると

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} (\boldsymbol{y}^T \boldsymbol{y} - 2\boldsymbol{y}^T X \boldsymbol{\beta} + \boldsymbol{\beta}^T X^T X \boldsymbol{\beta})$$

$$= \underbrace{-2(\boldsymbol{y}^T X)^T}_{\text{注2.1の(i)}} + \underbrace{2X^T X \boldsymbol{\beta}}_{\text{(same)の(ii)}}$$

$$= -2X^T \boldsymbol{y} + 2X^T X \boldsymbol{\beta}$$

よって正規方程式 $\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2X^T \boldsymbol{y} + 2X^T X \boldsymbol{\beta} = \boldsymbol{0}$

問2.3

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2 I_n)$$

$\boldsymbol{y}$ の従う分布は $N_n(X\boldsymbol{\beta}, \sigma^2 I_n)$. データ $\boldsymbol{y}, X$ が与えられたとき尤度関数は

$$L(\boldsymbol{\beta}, \sigma^2) \equiv f(\boldsymbol{y} \mid X; \boldsymbol{\beta}, \sigma^2)$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y} - X\boldsymbol{\beta})^T(\boldsymbol{y} - X\boldsymbol{\beta})\right\}$$

対数尤度関数は

$$\ell(\boldsymbol{\beta}, \sigma^2) \equiv \log L(\boldsymbol{\beta}, \sigma^2)$$

$$= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(\boldsymbol{y} - X\boldsymbol{\beta})^T(\boldsymbol{y} - X\boldsymbol{\beta})$$

$\boldsymbol{\beta}, \sigma^2$ の最尤推定値は $\ell(\boldsymbol{\beta}, \sigma^2)$ を最大化させる。このとき

$$\frac{\partial \ell(\boldsymbol{\beta}, \sigma^2)}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}}\left(\boldsymbol{y}^T\boldsymbol{y} - 2\boldsymbol{y}^T X\boldsymbol{\beta} + \boldsymbol{\beta}^T X^T X\boldsymbol{\beta}\right) \quad \leftarrow \text{(2.19)式}$$

$$= -2(\boldsymbol{y}^T X)^T + 2X^T X\boldsymbol{\beta}$$

$$= -2X^T\boldsymbol{y} + 2X^T X\boldsymbol{\beta}$$

$$= 0$$

かつ

$$\frac{\partial \ell(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(\boldsymbol{y} - X\boldsymbol{\beta})^T(\boldsymbol{y} - X\boldsymbol{\beta})$$

$$= 0$$

が成立。よってこれを解いて

$$\hat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \boldsymbol{y}, \quad \hat{\sigma}^2 = \frac{1}{n}(\boldsymbol{y} - X\hat{\boldsymbol{\beta}})^T(\boldsymbol{y} - X\hat{\boldsymbol{\beta}}) \blacksquare$$

問 2.4　結果が同じになる

問 2.5　$\mathbb{C} \in \mathbb{R}^{(p+1)\times 1}$, $A \in \mathbb{R}^{(p+1)\times(p+1)}$, $\beta = (\beta_0, \cdots, \beta_p)^T \in \mathbb{R}^{(p+1)\times 1}$

$$\frac{\partial(\mathbb{C}^T\beta)}{\partial\beta} = \frac{\partial}{\partial\beta}\left(\sum_{i=0}^{p} c_i \beta_i\right)$$
$$= \left(\frac{\partial(\sum_i c_i\beta_i)}{\partial\beta_0}, \cdots, \frac{\partial(\sum c_i\beta_i)}{\partial\beta}\right)^T$$
$$= (c_0, \cdots, c_p)^T$$
$$= \mathbb{C} \blacksquare$$

$$\frac{\partial(\beta^T A\beta)}{\partial\beta} = \frac{\partial}{\partial\beta}\left[(\beta_0, \cdots, \beta_p)\begin{pmatrix} a_{00} & \cdots & a_{0,p} \\ \vdots & \ddots & \vdots \\ a_{p,0} & \cdots & a_{p,p} \end{pmatrix}\begin{pmatrix}\beta_0 \\ \vdots \\ \beta_p\end{pmatrix}\right]$$
$$= \frac{\partial}{\partial\beta}\left[(\beta_0, \cdots, \beta_p)\begin{pmatrix} \sum_{i=0}^{p} a_{0,i}\beta_i \\ \vdots \\ \sum_{i=0}^{p} a_{p,i}\beta_i \end{pmatrix}\right]$$
$$= \frac{\partial}{\partial\beta}\left(\sum_{i=0}^{p} \beta_0 a_{0,i}\beta_i + \cdots + \sum_{i=0}^{p}\beta_p a_{p,i}\beta_i\right)$$
$$= \left(\sum_{i=0}^{p} a_{0,i}\beta_i + \sum_{i=0}^{p} a_{1,i}\beta_i + \cdots + \sum_{i=0}^{p} a_{p,i}\beta_i\right)$$
$$+ \left(\sum_{i=0}^{p} a_{i,0}\beta_i + \cdots + \sum_{i=0}^{p} a_{i,p}\beta_i\right)$$
$$= A\beta + A^T\beta = (A + A^T)\beta \blacksquare$$

Aが対称行列のとき $A = A^T$, 上の公式を使えば

$$\frac{\partial(\beta^T A\beta)}{\partial\beta} = (A + A^T)\beta$$
$$= 2A\beta \blacksquare$$

問 2.6

$\mathbb{C} \in \mathbb{R}^{m\times 1}$, $A \in \mathbb{R}^{m \times n}$, $Z = \mathbb{C} + AY$.

$$
\begin{aligned}
E[Z] &= E[\mathbb{C} + AY] \\
&= c + A\,E[Y] \quad \blacksquare
\end{aligned}
$$

$$
\begin{aligned}
E\left[(Z - E(Z))(Z - E(Z))^T\right] &= E\left[(\mathbb{C} + AY - (c + AE[Y]))\right. \\
&\qquad \left.(\mathbb{C} + AY - (c + AE[Y]))^T\right) \\
&= E\left[(A(Y - E[Y]))((A(Y - E[Y]))^T)\right] \\
&= A\,E\left[(Y - E[Y])(Y - E(Y))^T\right] A^T \\
&= A \operatorname{cov}(Y) A^T \quad \blacksquare
\end{aligned}
$$

問 2.7

$$
\begin{aligned}
P^2 &= \left(X(X^T X)^{-1} X^T\right)\left(X(X^T X)^{-1} X^T\right) \\
&= X(X^T X)^{-1} \underbrace{X^T X (X^T X^{-1})}_{\substack{\shortparallel \\ I_n}} X^T \\
&= X(X^T X)^{-1} X^T \\
&= P \quad \blacksquare
\end{aligned}
$$

$$
\begin{aligned}
(I_n - P)^2 &= \left(I_n - X(X^T X)^{-1} X^T\right)\left(I_n - X(X^T X)^{-1} X^T\right) \\
&= I_n - 2\underbrace{X(X^T X)^{-1} X^T}_{\substack{\shortparallel \\ P}} + \underbrace{\left(X(X^T X)^{-1} X^T\right)\left(X(X^T X)^{-1} X^T\right)}_{\substack{\shortparallel \\ P^2 = P}} \\
&= I_n - P \quad \blacksquare
\end{aligned}
$$

問2.8

$$\ell(\hat{\beta}, \hat{\sigma}^2) = -\frac{n}{2}\log(2\pi\hat{\sigma}^2) - \frac{1}{2\hat{\sigma}^2}\underbrace{(y - X\hat{\beta})^T(y - X\hat{\beta})}_{= n\hat{\sigma}^2}$$

$$= -\frac{n}{2}\log(2\pi\hat{\sigma}^2) - \frac{n}{2}$$

自由パラメータ数は $p+1$ コの回帰係数と 1 コの誤差分散の合計 $p+2$ コ

以上より

$$AIC = -2\left(-\frac{n}{2}\log(2\pi\hat{\sigma}^2) - \frac{n}{2}\right) + 2(p+2)$$

$$= n\log(2\pi\hat{\sigma}^2) + n + 2(p+2) \ \blacksquare$$